Available online at www.sciencedirect.com
ScienceDirect
Procedia Engineering 70 (2014) 85 – 92
Procedia Engineering
www.elsevier.com/locate/procedia

12th International Conference on Computing and Control for the Water Industry, CCWI2013

Detecting pipe bursts using Heuristic and CUSUM methods

M. Bakker[a,b,]*, D. Jung[c], J. Vreeburg[d,e], M. van de Roer[f], K. Lansey[c], L. Rietveld[a]

[a]*Delft University of Technology, P.O. Box 5048, 2600 GA Delft, The Netherlands*
[b]*Royal HaskoningDHV, P.O. Box 1132, 3800 BC, The Netherlands*
[c]*The University of Arizona, P.O. Box 210073, AZ 85721-0073, Tucson, Arizona, United States of America*
[d]*Wageningen University, P.O. Box 17, 6700 AA Wageningen, The Netherlands*
[e]*KWR Watercycle Research Institute, P.O. Box 1072, 4330 BB Nieuwegein, The Netherlands*
[f]*Dunea Duin & Water, P.O. Box 756, 2700 AT, Zoetermeer, The Netherlands*

Abstract

Pipe bursts in a drinking water distribution system lead to water losses, interruption of supply, and damage to streets and houses due to the uncontrolled water flow. To minimize the negative consequences of pipe bursts, an early detection is necessary. This paper describes a heuristic burst detection method, which continuously compares forecasted and measured values of the water demand. The forecasts of the water demand were generated by an adaptive water demand forecasting model. To test the method, a dataset of five years of water demand data in a supply area in the Western part of the Netherlands was collected. The method was tested on a subset of the data (only the winter months) in which 9 (larger) burst events were reported. The detection probability for the reported bursts was 44.4%, at an acceptable rate of false alarms of 5.0%. The results were compared with the CUSUM method, which is a general statistical process control (SPC) method to identify anomalies in time series. The heuristic and CUSUM methods generated comparable results, although rate of false alarm for the heuristic method was lower at the same detection probability.

Selection and peer-review under responsibility of the CCWI2013 Committee

Keywords: Pipe burst detection; Demand forecasting, SPC methods.

* Corresponding author. Tel.: +31 88 348 2327; fax: +31 88 348 2801.
E-mail address: martijn.bakker@rhdhv.com

Selection and peer-review under responsibility of the CCWI2013 Committee
doi:10.1016/j.proeng.2014.02.011

1. Introduction

1.1. Pipe bursts in drinking water distribution systems

Pipe bursts and leakages are part of the normal day to day operation of water supply companies. The burst rate in the Netherlands is on average 0.07-0.09 bursts per km of water main per year (Trietsch and Vreeburg, 2005). For the whole country this results in 10,000 bursts per year (28 bursts per day). This means that an average water supply company in the Netherlands needs to respond to 2-3 bursts per day. As the majority of the total length of the water distribution network consists of pipes with a diameter of 100 mm or smaller, the majority of the bursts also occurs on these smaller pipes. Thornton et al. (2008) classified leakages in water distribution networks in ‘background’ (small continuous running leakages), ‘unreported’ (slightly bigger leakages that tend to increase and need attention) and ‘reported’ (big leaks that need to be repaired as soon as possible). After the occurrence of a burst of the latter category, it takes some time before the burst is actually reported to the utility, and the utility is aware of the situation. In the time frame between the occurrence of the burst and the point in time that the broken pipe is isolated, the burst causes negative consequences like water loss, interruption of supply, and damage to streets and houses due to the uncontrolled water flows.

To minimize the negative effects of pipe bursts, early identification is critical. Research to develop a heuristic detection method for pipe bursts is presented here. The method is based on monitoring the difference between the measured and forecasted water demand in a supply area. The results of the heuristic method are compared to the CUSUM method. CUSUM is a statistical process control (SPC) method, that is commonly used for anomaly detection.

1.2. Previous work

For the detection of pipe bursts, various techniques can be used. Puust et al. (2010) give a profound overview of different techniques for managing background leakage in distribution systems, as well as detecting pipe bursts.

Monitoring pressure transients

One of the commonly used techniques for pipe burst detection is based on monitoring pressure transients in the distribution system, which occur after a sudden failure (rupture) of a pipe. By measuring pressure at different locations at a very high sampling rate (2000 Hz, Misiunas et al. (2005a)) the propagation of the pressure transient in the network can be measured, and the burst location can be approximated. Colombo et al. (2009) present a literature overview of transient monitoring techniques. Brunone and Ferrante (2001), Misiunas et al. (2005a) Misiunas et al. (2005b), Kim (2005), Duan et al. (2011), and Kwon and Lee (2011) present theoretical research to further develop this technique. The technique is only applicable for actual bursts. Pipe failure which develops gradually will not induce a pressure transient, and will therefore not be detected by this technique.

Monitoring flow, or pressure and flow in a DMA

When flow and pressure measurements are available for off-line district metered area (DMA) monitoring, the measurements can be used for on-line monitoring when made available (semi) online. This technique was researched and tested in a real water supply system in North Yorkshire, UK (Mounce et al., 2002; Mounce et al., 2003; Mounce and Machell, 2006; Mounce and Boxall, 2010; Mounce et al., 2011). The papers describe the application of artificial neural networks combined with fuzzy logic to evaluate pressure and flow measurements. In Mounce et al. (2011) the application of the system in practice in a six month test period, is described. It was proved that the system was able to detect 7 out of 18 reported bursts (11 events missed), where the system generated a total of 46 alerts (39 were not related to actual bursts).

Other promising research in the field of burst detection was carried out by Palau et al. (2011), who used a multivariable statistical technique (principle component analysis) to derive burst events from flow and pressure data. Bicik et al. (2011) combined flow and pressure data with information from other data sources, like customer contacts and a hydraulic model to detect burst events. Where Khan et al. (2005) describe the application of experimental failure sensors measuring opacity or temperature for the detection of pipe bursts.

2. Materials and methods

2.1. Study area and dataset

In order to develop and test the heuristic and CUSUM methods to detect pipe burst, a dataset with historic flows was collected. For a supply area of the water utility Dunea in the Western part of the Netherlands, all measured flows with 5 minutes intervals of the period 2008-2012 were collected (526,176 values per time series). The flows were measured at permanent assets of the water utility (treatments plant, reservoirs, pumping stations, permanent measuring points) which were all connected to the central automation network. All data was stored in a central database system. Due to the high reliability of the meters, the automation network, and the database system, virtually no data gaps or data errors were present in the datasets. Also information was available of all repaired pipe burst / leakage incidents in the area in the same period. To be able to develop and test the method, we chose to use only a subset of the complete dataset, where demands were rather stable. Therefore a subset was created, which only contained data of the winter months January, February and December. The researched area is shown in Fig. 1 and the characteristics and reported incidents in the winter months are summed in Table 1.

Table 1. Characteristics of the researched area (values of 2008-2012), and the selected subset of the data.

Area	# connections	water demand (m^3/h)	water use (m^3/conn./year)	# days with undisturbed supply	# reported burst incidents
1. Rhine area	130,920	2,290	145	359	9

The area contains 130,920 connections, which is a very large number compared to district metered areas (DMAs) in other countries, that generally contain 1,000-3,000 connections (Thornton et al., 2008). The water supplied in this area, is mainly produced at the Katwijk Water Treatment Plant (1.), and buffered in the clear water reservoirs Cronestein (2.), Noordwijkerhout (3.), and De Engel (4.). These reservoirs are filled during low demand (at night) with water from the network by a flow controlled valve, and water is pumped back to the network during demand peaks by flow controlled or on/off pumps. The net water demand in the area was determined by performing a water balance calculation.

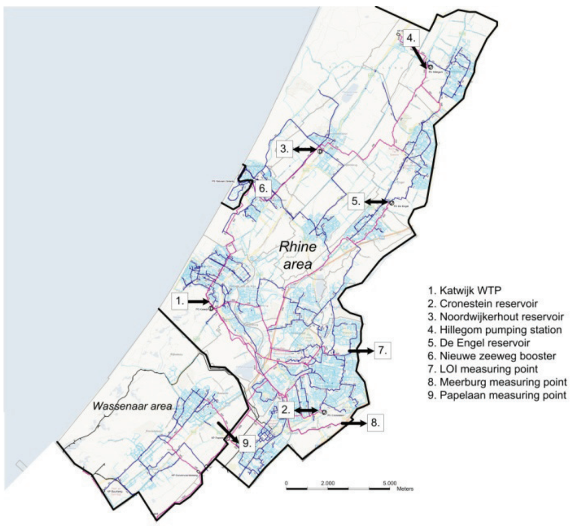

Fig. 1. Area of the case study. The net water demand in the area was calculated using the flow measurements 1. to 9.

2.2. Heuristic burst detection method

We developed a heuristic burst detection method that consists of three main steps: 1. Generating forecasted values of the water demand; 2. Transformation of the forecasted and measured values; 3 Analysis of the differences between forecasted and measured values, for A: generating control limits based on historic data, and B: generating alerts and alarms, using real-time data.

Forecasting demand

The net water demand in the area was the input for the water demand forecasting model. In the heuristic burst detection method, the adaptive water demand forecasting model described by Bakker et al. (2013b) was used. This model generates a water demand forecast for the next 48 hours with 15-minutes time steps. The model adaptively learns the average demand patterns and factors for the seven days of the week, and for a configurable number of deviant day types (like national holidays and primary school holiday periods). The model was originally developed for optimal control (Bakker et al., 2013a), but was in this research used for detection of pipe bursts. For detection of pipe bursts, only the actual forecasted value (the so-called nowcast) was used, rather than all forecasted values for the next 48 hours. Fig. 2 shows two examples of trends of measured and forecasted water.

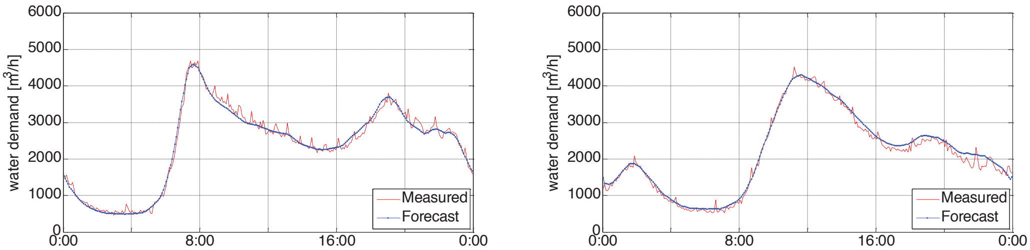

Fig. 2. Examples of trends of measured and forecasted water demand. (a) a normal weekday; (b) a deviant day (New Years Day).

Signal transformation

The measured time series of water demand are somewhat noisy due to random temporal and spatial variation in the water demand. The noise in the signal can be reduced by transforming the signal to a moving average over different time frames, where the noise decreases as the time frame is increased. Reducing the noise enables closer monitoring without increasing the number of false alarms. However, when calculating the moving average, the deviant values shortly after the burst will be leveled out by normal values prior to the burst. Only when the burst (and thus the deviant demands) persist longer, the burst can be detected. This means that by taking the moving averaged value over longer time frames, smaller bursts can be detected but only some time after the burst occurred. The heuristic method for detecting pipe burst includes monitoring the moving averaged signal over time frames of 5, 10, 15, 30, 60, 120 and 240 minutes.

Forecast errors analysis

Differences between measured and forecasted (transformed) values indicate a pipe burst. For monitoring purposes, control limits need to be set to distinguish between normal forecasting inaccuracies on the one hand and possible or certain burst events on the other hand. To determine the control limits, we analyzed the forecasting errors in the year prior to the monitoring year. When analyzing the water demand forecast errors, we did two observations: 1. the errors were not normally distributed around the average error, but the relative larger underestimates and overestimates have a higher probability than may be expected based on the standard deviation, and 2. the relative error (compared to the forecasted value) decreased as the forecasted value increased. Based on these observations, the control limit was set to the 5% exceedance probability of the forecast error, which proved to be a more generally applicable measure than the standard deviation. And the forecasted values were divided in five classes (from low to high demand) for which different 5% exceedance probability values were derived. Fig. 3

shows an example of the analysis of the forecast errors, and the resulting 5% exceedance probability values which were used to derive the control limits.

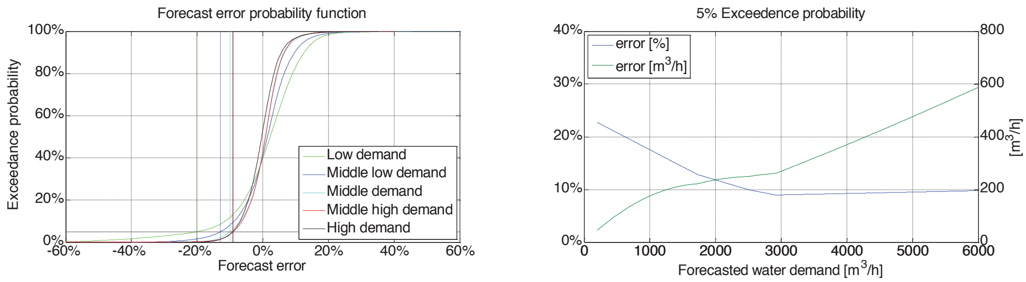

Fig. 3. Example of the forecast error analysis of the five minutes time frame water demand forecast of the Rhine area. (a) forecast error probability functions of the five classes of forecasted water demand; (b) 5% exceedance probability function.

When monitoring the water distribution network, the heuristic burst detection method compares the actual forecasting errors of all signals (non-transformed and transformed water demands) with the control limits which are based on the 5% exceedance probability. An alarm is generated if the error is larger than the 5% exceedance probability value multiplied by *n*. The *n*-factor determines how sensitive the flow is being monitored: a high *n*-factor results in high control limits, which means that only larger bursts will be detected at a relative low number of false alarms; with a lower *n*-factor, smaller bursts can be detected but at higher number of false alarms. In this study, we used an *n*-factor of 3.0 for "tight" monitoring, and 4.5 for "loose" monitoring.

2.3. Cumulative sum (CUSUM) control chart

The cumulative sum (CUSUM) method is a general statistical process control (SPC) method for identification of anomalies. Misiunas et al. (2005a) applied the CUSUM method for burst detection and location. The method utilizes information from a sequence of measured values by calculating the cumulative sums of the deviation of the measured values from the mean value:

$$C_i^+ = \max[0, x_i - (\mu_i + K_u) + C_{i-1}^+] \tag{1}$$

$$C_i^- = \max[0, (\mu_i - K_u) - x_i + C_{i-1}^-] \tag{2}$$

where C_i^+ and C_i^- are one-side upper and lower cumulative sum up to and including the i^{th} time period and $C_0^+ = C_0^- = 0$; x_i is the monitored quality characteristics values (in this case the net water demand) at the i^{th} time period; μ_i is the mean value of the quality characteristics' value at the i^{th} time period; K_u is a reference value. If either C_i^+ or C_i^- exceeds the decision interval H, the process is considered to be out of control. The mean value μ_i of net water demand for every 5 minutes was calculated from the 359 normal days of the dataset, where a set of mean values were derived for normal weekdays, and a set of values for weekend days and other deviant days.

The two parameters, K_u and H, were estimated to maximize the detection efficiency. Like in the heuristic method, the parameters can be set for "tight" and for "loose" monitoring. In this study, we used K_u=2.6 and H=1 for "tight" monitoring, and K_u=2.9 and H=4 for "loose" monitoring.

3. Results

Simulations with the historic dataset were carried out to test the burst detection methods. An example of a detected and a not detected burst incident with the heuristic method are shown in Fig. 4. The figure also shows varying control limits, depending on the forecasted demand (the limit decreases as the forecasted value decreases) and the moving average time frame (the limit decreases as the length of the time increases).

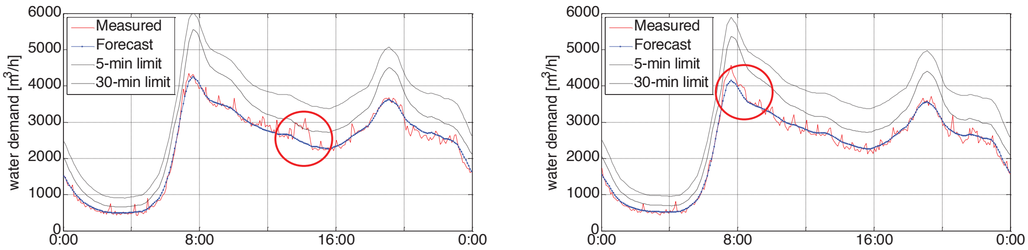

Fig. 4. Example of measured and forecasted demand, including the 5-minutes and 30-minutes Moving Average control limits of the heuristic method. (a) detected burst incident (burst 5, at 14:00); (b) not detected burst incident (burst 2, at 7:15).

To evaluate the researched detection methods, the detection probability (*DP*) and the rate of false alarm (*RF*) as proposed by Jung et al. (2013) were calculated, with:

$$DP = \frac{Number\ of\ detected\ burst}{Total\ number\ of\ bursts} \cdot 100\% \qquad (3)$$

$$RF = \frac{Number\ of\ \ false\ alarms}{Number\ of\ evaluated\ days\ without\ incident} \cdot 100\% \qquad (4)$$

In the researched dataset the number of evaluated days was 359, and the number of (relative larger) burst events was 9. The results of the evaluated detection methods on this dataset are shown in Table 2.

Table 2. Results heuristic and CUSUM burst detection method.

			"Loose" monitoring		"Tight" monitoring	
			Heuristic	CUSUM	Heuristic	CUSUM
Detection Probability (*DP*)			44.4%	44.4%	66.7%	66.7%
Rate of False Alarm (*RF*)			5.0%	6.1%	12.5%	26.5%
Burst 1: 2 February 2008	08:00	(± 150 m^3/h – 6% avg. flow)	not detected	not detected	not detected	not detected
Burst 2: 4 February 2009	07:15	(± 180 m^3/h – 8% avg. flow)	not detected	not detected	not detected	detected
Burst 3: 20 February 2009	09:00	(± 100 m^3/h – 4% avg. flow)	not detected	not detected	not detected	not detected
Burst 4: 21 February 2009	04:45	(± 200 m^3/h – 9% avg. flow)	detected	detected	detected	detected
Burst 5: 3 December 2009	12:30	(± 400 m^3/h – 17% avg. flow)	not detected	not detected	detected	not detected
Burst 6: 9 December 2010	22:30	(± 500 m^3/h – 22% avg. flow)	detected	detected	detected	detected
Burst 7: 4 January 2011	14:30	(± 1900 m^3/h – 83% avg. flow)	detected	detected	detected	detected
Burst 8: 13 December 2011	14:40	(± 1200 m^3/h – 52% avg. flow)	detected	detected	detected	detected
Burst 9: 18 January 2012	13:40	(± 400 m^3/h – 17% avg. flow)	not detected	not detected	detected	detected

Table 2 shows that bursts 1. to 3. were relatively small bursts (4 to 8% of the average water demand) that occurred in the morning peak demand (between 7:00 and 9:00). The deviant flow rate caused by these bursts was too small to be detected at that point in time, and therefore the bursts were not detected by any of the methods. When applying "loose" monitoring, bursts 5 and 9 could not be detected, despite the fact that the flow rates of these bursts were not too low (17% of the average demand). The reason for this is that the bursts occured at a point in time when the demand was rather high, so the percentage burst flow rate at that time was rather low. And a second aspect is that both bursts only lasted for 1.5 hour. The relative low flow rate combined with a short duration made the bursts more difficult to detect. These bursts could only be detected when the parameters were set to monitor more precisely ("tight" monitoring). As disadvantage of the tight monitoring is that the rate of false alarm (*RF*) increased largely.

Table 2 also shows that the same *DP* could be achieved with heuristic and the CUSUM method. The difference between the methods is that the heuristic method was able to detect the bursts at a lower *RF*. Especially when "tight" monitoring is applied, the *RF* of the heuristic method was significantly lower.

4. Discussion and conclusions

We developed a heuristic burst detection method and compared this method with the general CUSUM method.

Discussion

The reason for the lower *RF* of heuristic method compared to the CUSUM method, is that the heuristic method is quite specific for the analyzed dataset. Quite some effort was put in optimizing the forecasting model used in the heuristic method in order to generate accurate water demand forecasts. When comparing the accurate forecasts with the measured demand, bursts could be detected at a rather low *RF*. The CUSUM method on the other hand, uses only two water demand patterns to compare the measured water demand with. This makes the CUSUM method more generally applicable for anomaly detection, but less accurate compared to the more specific heuristic method.

Conclusions

The heuristic burst detection method, based on monitoring the net water demand in an area, was able to detect all pipe bursts where the burst flow exceeds 20-25% of the average flow. Smaller pipe burst could only be detected if they occurred when the demand was low (at night) or when a higher *RF* is accepted. Comparable detection results could be achieved when monitoring the net demand in an area with a CUSUM method, although the *RF* was higher compared to the heuristic model.

Monitoring the net water demand in area can contribute to detect pipe bursts at an early stage. An early detection is essential to minimize the negative aspects of pipe bursts, and can therefore contribute to the resilience of water supply systems.

Acknowledgements

This study was carried out in the DisConTO project (Distribution Control Training & Operation). The project is a cooperation between 4 water supply companies (Vitens, Dunea, PWN and Brabant Water), Delft University of Technology, The National Institute for Public Health and the Environment (RIVM), Royal HaskoningDHV Consultancy and Engineering and UReason. The project is financially supported by the Dutch government through the "Innowator" program.

References

Bakker, M., Vreeburg, J.H.G., Palmen, L.J., Sperber, V., Bakker, G., and Rietveld, L.C. (2013a). "Better water quality and higher energy efficiency by using model predictive flow control at water supply systems". *Journal of Water Supply: Research and Technology - AQUA.* 62 (1): 1-13.

Bakker, M., Vreeburg, J.H.G., Van Schagen, K.M., and Rietveld, L.C. (2013b). (submitted). "A fully adaptive forecasting model for short-term drinking water demand ". *Journal of Environmental Modelling and Software.*

Bicik, J., Kapelan, Z., Makropoulos, C., and Drasavić, G.A. (2011). "Pipe burst diagnostics using evidence theory". *Journal of Hydroinformatics.* 13 (4): 596-608.

Brunone, B., and Ferrante, M. (2001). "Detecting leaks in pressurised pipes by means of transients". *Journal of Hydraulic Research.* 39 (5): 539-547.

Colombo, A.F., Lee, P., and Karney, B.W. (2009). "A selective literature review of transient-based leak detection methods". *Journal of Hydro-Environment Research.* 2 (4): 212-227.

Duan, H.F., Lee, P.J., Ghidaoui, M.S., and Tung, Y. (2011). "System response function-based leak detection in viscoelastic pipelines". *Journal of Hydraulic Engineering.* 138 (2): 143-153.

Jung, D., Kang, D., Liu, J., and Lansey, K.E. (2013). "Improving resilience of water distribution system through burst detection", *Proceedings 2013 ASCE Environment and Water Resources Institue Annual Conference*, Cincinnati (OH) United States of America.

Khan, A., Widdop, P.D., Day, A.J., Wood, A.S., Mounce, S.R., and Machell, J. (2005). "Performance assessment of leak detection failure sensors used in a water distribution system". *Journal of Water Supply: Research and Technology - AQUA.* 54 (1): 25-36.

Kim, S.H. (2005). "Extensive development of leak detection algorithm by impulse response method". *Journal of Hydraulic Engineering.* 131 (3): 201-208.

Kwon, H.J., and Lee, C.E. (2011). "Probability of pipe breakage regarding transient flow in a small pipe network". *Annals of Nuclear Energy.* 38 (2-3): 558-563.

Misiunas, D., Lambert, M., Simpson, A., and Olsson, G. (2005a). "Burst detection and location in water distribution networks". *Water Science and Technology: Water Supply.* 5 (3-4): 71-80.

Misiunas, D., Vítkovský, J., Olsson, G., Simpson, A., and Lambert, M. (2005b). "Pipeline break detection using pressure transient monitoring". *Journal of Water Resources Planning and Management.* 131 (4): 316-325.

Mounce, S.R., Day, A.J., Wood, A.S., Khan, A., Widdop, P.D., and Machell, J. (2002). "A neural network approach to burst detection". *Water Science and Technology.* 45 (4-5): 237-246.

Mounce, S.R., Khan, A., Wood, A.S., Day, A.J., Widdop, P.D., and Machell, J. (2003). "Sensor-fusion of hydraulic data for burst detection and location in a treated water distribution system". *Information Fusion.* 4 (3): 217-229.

Mounce, S.R., and Machell, J. (2006). "Burst detection using hydraulic data from water distribution systems with artificial neural networks". *Urban Water Journal.* 3 (1): 21-31.

Mounce, S.R., and Boxall, J.B. (2010). "Implementation of an on-line artificial intelligence district meter area flow meter data analysis system for abnormality detection: A case study". *Water Science and Technology: Water Supply.* 10 (3): 437-444.

Mounce, S.R., Mounce, R.B., and Boxall, J.B. (2011). "Novelty detection for time series data analysis in water distribution systems using support vector machines". *Journal of Hydroinformatics.* 13 (4): 672-686.

Palau, C.V., Arregui, F.J., and Carlos, M. (2011). "Burst detection in water networks using principal component analysis". *Journal of Water Resources Planning and Management.* 138 (1): 47-54.

Puust, R., Kapelan, Z., Savic, D.A., and Koppel, T. (2010). "A review of methods for leakage management in pipe networks". *Urban Water Journal.* 7 (1): 25-45.

Thornton, J., Sturm, R., and Kunkel, G. 2008. *Water Loss Control*. 2nd ed: McGraw-Hill, New York, USA.

Trietsch, E.A., and Vreeburg, J.H.G. (2005). "Reliability of valves and section isolation". *Water Science and Technology: Water Supply.* 5 (2): 47-51.